Classification of Interest Rate Curves Using Self-Organising Maps

M.Kanevski [a,*], M.Maignan [b], V.Timonin [a,1], A.Pozdnoukhov [a,1]

[a]*Institute of Geomatics and Analysis of Risk (IGAR), Amphipole building, University of Lausanne, CH 1015 Lausanne, Switzerland*

[b]*Banque Cantonale de Genève (BCGE), Geneva, Switzerland*

Abstract

The present study deals with the analysis and classification of interest rate curves. Interest rate curves (IRC) are the basic financial curves in many different fields of economics and finance. They are extremely important tools in banking and financial risk management problems. Interest rates depend on time and maturity which defines term structure of the interest rate curves. IRC are composed of interest rates at different maturities (usually fixed number) which move coherently in time. In the present study machine learning algorithms, namely Self-Organising maps - SOM (Kohonen maps), are used to find clusters and to classify Swiss franc (CHF) interest rate curves.

Key words: interest rates curves, clustering, self-organising maps
PACS: 89.65.Gh, 05.45.Tp

1 Introduction

Interest rates depend on time and maturity which define term structure of the interest rate curves. The temporal evolution of interest rates data for all maturities considered in this study is presented in Figure 1. Usually, this information is available for several fixed time intervals (daily, weekly, monthly) and for some definite maturities. In this study interest rate curves are composed of LIBOR rates with maturities 1 week, 1, 2, 3, 6, 9, 12 months and

---

* Corresponding author.
  Email address: `Mikhail.Kanevski@unil.ch` (M.Kanevski).

Fig. 1. Temporal evolution of CHF interest rates during years 1999-2006.

	1W	1M	2M	3M	6M	9M	12M	2_YEAR	3_YEAR	4_YEAR	5_YEAR	7_YEAR	10_YEAR
1W	1.00	0.99	0.98	0.98	0.96	0.95	0.93	0.88	0.84	0.82	0.79	0.74	0.68
1M	0.99	1.00	1.00	0.99	0.98	0.96	0.95	0.91	0.87	0.84	0.82	0.77	0.71
2M	0.98	1.00	1.00	1.00	0.99	0.98	0.97	0.92	0.89	0.86	0.84	0.79	0.73
3M	0.98	0.99	1.00	1.00	0.99	0.99	0.98	0.94	0.91	0.88	0.86	0.81	0.75
6M	0.96	0.98	0.99	0.99	1.00	1.00	0.99	0.96	0.93	0.91	0.89	0.84	0.78
9M	0.95	0.96	0.98	0.99	1.00	1.00	1.00	0.97	0.95	0.93	0.90	0.86	0.80
12M	0.93	0.95	0.97	0.98	0.99	1.00	1.00	0.98	0.96	0.94	0.92	0.88	0.82
2_YEAR	0.88	0.91	0.92	0.94	0.96	0.97	0.98	1.00	0.99	0.98	0.97	0.94	0.89
3_YEAR	0.84	0.87	0.89	0.91	0.93	0.95	0.96	0.99	1.00	1.00	0.99	0.96	0.92
4_YEAR	0.82	0.84	0.86	0.88	0.91	0.93	0.94	0.98	1.00	1.00	1.00	0.98	0.95
5_YEAR	0.79	0.82	0.84	0.86	0.89	0.90	0.92	0.97	0.99	1.00	1.00	0.99	0.96
7_YEAR	0.74	0.77	0.79	0.81	0.84	0.86	0.88	0.94	0.96	0.98	0.99	1.00	0.99
10_YEAR	0.69	0.71	0.73	0.75	0.78	0.80	0.82	0.89	0.92	0.95	0.96	0.99	1.00

Fig. 2. Correlation matrix for all maturities.

swap rates with maturities 1, 2, 3, 4, 5, 7 and 10 years. The behaviour of different maturities is coherent and consistent in time which can be confirmed by the corresponding global correlation matrix between different maturities (Figure 2).

There are some well known important stylised facts (typical and stable behaviour) that have to be taken into account during analysis, modelling and interpretation of IRC (Diebold and Canlin, 2006):

- The average yield curve is increasing and concave.
- The yield curve assumes a variety of shapes through time, including upward sloping, downward sloping, humped, and inverted humped.
- IR dynamics is persistent, and spread dynamics is much less persistent.
- The short end of curve is more volatile than the long end.
- Long rates are more persistent than short rates.

The main task of the present study is the analysis of IRC as the objects embedded into 13-dimensional space (equal to the number of maturities) and application of nonparametric nonlinear tool - Self-Organising maps - in order to find classes of typical behaviour of IRC. A priori hypothesis which will

2

be verified in this study is that the curves are clustered in time reflecting the market conditions. Another fact supporting the idea that a few typical IRC exist is a parametric Nelson-Siegel model. In this model IRC are parameterised by 3 factors corresponding to long-term, short-term and medium term IR behaviour. These parameters can be interpreted in terms of level, slope (maturity$_{10years}$ - maturity$_{3months}$) and curvature ($2 \cdot$ maturity$_{2years}$ - maturity$_{3months}$ - maturity$_{10years}$) of the curves. This parametric approach was used to forecast IRC through forecasting these 3 parameters using linear ARMA type models (Diebold and Canlin, 2006).

2 Self-Organising Maps

Self-Organising maps (SOM) belong to the unsupervised machine learning algorithms. The unsupervised learning methods solve clustering, classification and density modelling problems using unlabeled data. SOM are widely used for the dimensionality reduction and high-dimensional data visualisation (projection onto two-dimensional space). Unlabeled data are points/vectors in a high-dimensional feature space that have some attributes (or coordinates) but have no target values, neither continuous (regression) nor categorical (classification). The main task of SOM is to "group" or to "range" these input vectors according to the defined similarity measure and to catch regularities (or patterns) in the data preserving its topological structure. The detailed presentation of SOM along with a comprehensive review of their application, including socio-economic and financial data is given in (Kohonen T., 2000).

2.1 The structure and initialisation of SOM

Self-organising map can be represented as a single layer feedforward network where the output neurons are arranged in a two-dimensional topological grid. The grid can be eihther rectangular or hexagonal. In the first case each neuron (except borders and corners) has four nearest neighbours, in the second - six ones. The hexagonal map requires more calculations and provides more smoothed result. Attached to every neuron there is a weight vector with the same dimensionality as the input space. Each unit $\boldsymbol{i}$ has a corresponding weight vector $\boldsymbol{w}_i = \{\boldsymbol{w}_{i1}, \boldsymbol{w}_{i2}, \ldots, \boldsymbol{w}_{id}\}$ where d is a dimension of the input feature space. The SOM learning procedure consists in finding the parameters of the network (weights $\boldsymbol{w}$) in order to group the data into clusters keeping its topological structure, thus finding a projection of high-dimensional data into a low-dimensional space.

Competitive learning is used for training the SOM, i.e. output neurons compete

among themselves to share the input data samples. The winning neuron $\boldsymbol{w}_w$ is a neuron which is the "closest" to the input example $\boldsymbol{x}$ among all other m neurons in the defined metric:

$$d(x, w_w) = \min_{1 \leq j \leq m} d(x, w_j). \tag{1}$$

The first step of the SOM learning is the initialisation of the neurons' weights. Two methods are widely used for the latter (Kohonen T., 2000). Initial weights can be taken as the coordinates of randomly selected m points from the data set. Otherwise, small random values can be sampled evenly from the input data subspace spanned by the two largest principal component eigenvectors. The second method can increase the speed of training but may lead to a local minima and miss some non-linear structures in data.

2.2 Training procedure

The iterative training process for each i-th neuron is

$$w_i(t + 1) = w_i(t) + h_i(t)[x(t) - w_i(t)] \tag{2}$$

where $h_i(t)$ is a so-called *neighbourhood function*. It is defined as a function of time t (or, more precisely, a training iteration) and defines the neighbourhood area of the i-th neuron.

The simplest neighbourhood function refers to a neighbourhood set of array points around the node i. Let this index set be denoted as R, whereby

$$\begin{cases} h_i(t) = \alpha(t), & \text{if } i \in Rand \\ h_i(t) = 0, & \text{if } i \notin R \end{cases} \tag{3}$$

where $\alpha(t)$ is a *learning rate* defined by some monotonically decreasing function of time, such that $0 < \alpha(t) < 1$.

Another widely used neighbourhood function is a Gaussian

$$h_i(t) = \alpha(t) \exp\left(-\frac{d(i, w)}{2\sigma^2(t)}\right). \tag{4}$$

The width of the kernel $\sigma(t)$ (corresponding to R) is a monotonically decreasing function of time as well. The exact forms of $\alpha(t)$ and $\sigma(t)$ are not
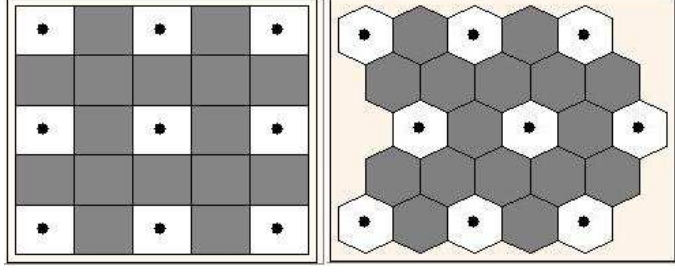
Fig. 3. Structure of the U-matrix for 3×3 rectangular (left) and hexagonal (right) SOM.

of particular importance. They should monotonically decrease in time. They tend to zero for $t \to T$, where T is the total number of iterations. For example, $\alpha(t) = c\ ((1\text{-}t)/T)$,where c is a predefined constant (for instance, $T/100$).

As a result, at the early stage of training when the neighbourhood is broad (and covers all the neurons), the self-organisation takes place at the global scale. At the end of training, the neighbourhood shrunks to zero and only one neuron updates its weights.

2.3 SOM visualisation tools

Several visualisation tools ("maps") are used to represent the trained SOM and to apply it for data analysis. **Hits** map shows how many times each neuron wins (the number of hits); **U-matrix** (**u**nified distance **matrix**) is a map of distances between each of neurons and all his neighbours (see details below); **Slices** are the 2D slices of the SOM weights (total number of maps is equal to the dimension of the data); **Clusters** is a map of recognised clusters in the SOM. All these maps are straightforward in the construction and interpretation except the *U-matrix*, which is described in more details below.

Dimension of the U-matrix for the $[xdim \ \times \ ydim]$ SOM is $[2 \ \cdot \ xdim - 1] \times [2 \ \cdot \ ydim - 1]$. Figure 3 presents an example of the U-matrix for 3×3 rectangular and hexagonal SOMs. The dimension of this U-matrix is 25. At the first step, the average (or median) distances between the nearest SOM neurons are calculated. These values are assigned to gray cells of the U-matrix (Figure 3). Secondly, the average (or median) of values of the nearest U-matrix cells, computed in step 1, are calculated and assigned to the white cells. For better visual perception, next step can be applied optionally. One treat the calculated U-matrix map as an image and smooth it out with a simple average (or median) filter, well-known in image processing. Finally, the resulting map is painted using some colour scale. In this study, light colour corresponds to larger distances between neighbouring nodes and thus indicates borders of clusters in the maps.
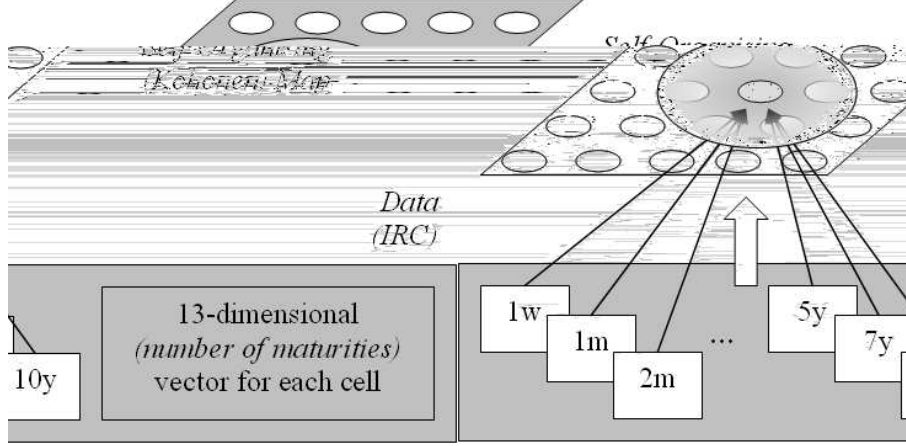
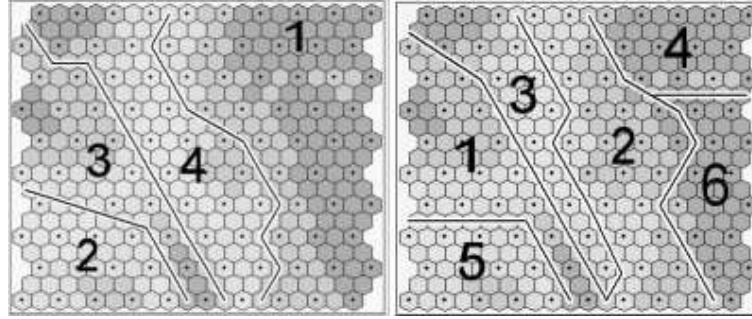Fig. 4. SOM structure for IRC classification.

Fig. 5. SOM map classified by k-Means clustering method. Predefined number of clusters is 4 (left) and 6 (right).

3 SOM classification of interest rate curves

SOM structure for the current study is presented in Figure 4. A 13-dimensional (corresponding to the number of maturities) input feature space was built. Only daily interest rate values were used for constructing the SOM maps. No temporal information was presented to the model. U-matrix of the trained SOM is used as a visualisation tool describing the structure of data. To detect the cluster structure of the data, K-Means algorithm was applied to group the cells of the U-matrix of the trained SOM.

Let us consider the results. It was found that 4 classes reveal the major IRC groups, while 3 classes are insufficient. The U-matrix of the trained SOM with 4 predefined clusters is presented in Figure 5. In Figure 6 the examples of the IR curves (rate vs. maturity) for each of 4 clusters are presented. Figure 7 presents the temporal structure (distribution of IRC classes in time) for each of 4 classes. One can see that the behaviour of the curves is similar inside the clusters and dissimilar for the different clusters. Let us examine the U-matrix
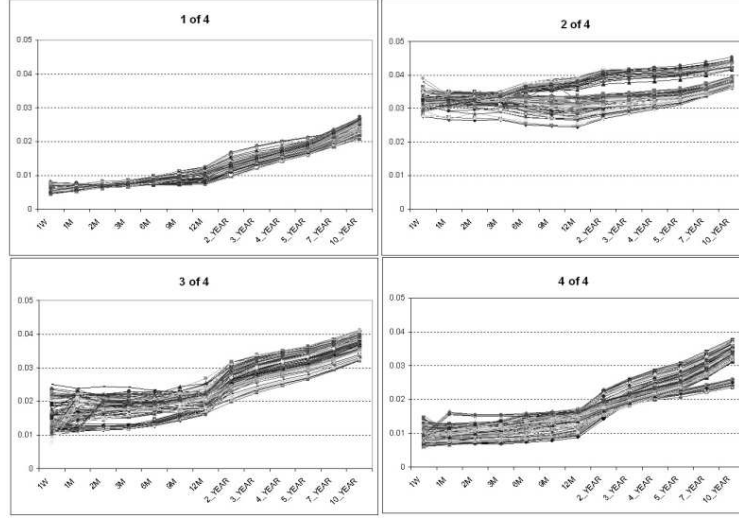
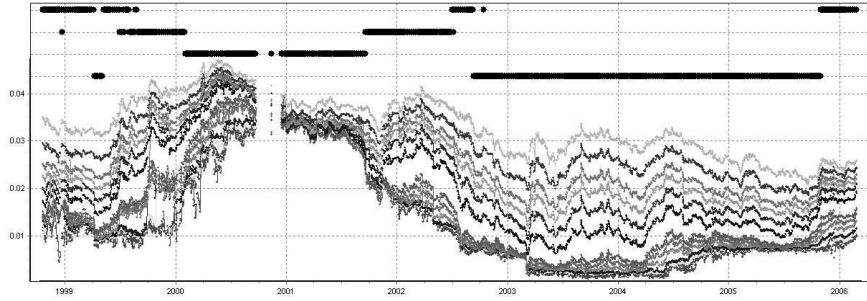Fig. 6. Examples of the IR curves (rate vs. maturity) for each of 4 clusters.

Fig. 7. Classification result for all maturities in time graphs. Black points on the top correspond to the cluster membership (marked at the right vertical axis).

(Figure 5) in details. The advantage of SOM mapping is that one can not only divide the data into clusters but it is also possible to find out more particular details in the structure of the data. For example, let us compare cluster 1 with cluster 2. Class 1 is located in the area where dark colours (dark colour corresponds to smaller distances between cells) dominate. On the contrary, class 2 is located in the area where light colours dominate. It means that IR curves from class 1 are more similar to each other than the IR curves from the class 2 (see Figures 6 and 7). Another useful feature of the SOM mapping is a presentation of a topological structure of the data. Fore example, it means that class 3 is closer (more similar) to the class 4 than to the class 2 (see Figures 6 and 7). More detailed structure of the IRC can be elaborated when considering 6 basic classes (Figures 5 (right) and 8). Note that in this case the former class 1 from 4-classes division was split into two (4 and 6) in the 6-classes division. The obtained class 3 (in 6-classes division) delineate the transition phase between former classes 3 and 4 (4-classes division).
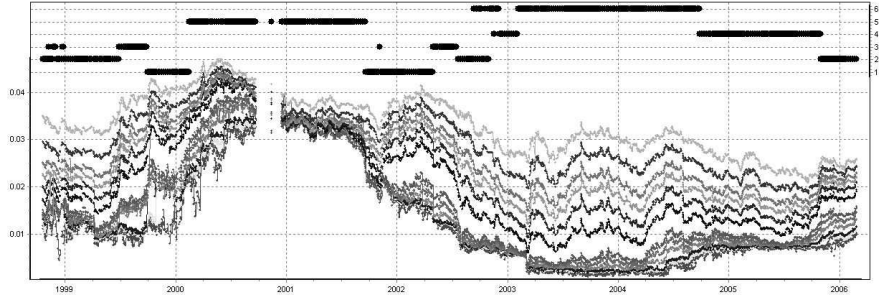
7

Fig. 8. Classification result for all maturities in time graphs.

4 Conclusions

Self-Organising Kohonen maps were applied to study interest rate curves clustering in time. An interesting finding is the revealed existence of several typical behaviours of curves and their clustering in time around low level rates, high level rates, and periods of transition between the two. Such analysis can help in the prediction of interest rate curves, evaluation of financial products and in financial risk management. Future studies dealing with IRC classification and market conditions are in progress.

References

Kohonen T. (2000). Self-organising maps. 3rd edition. Springer. 521 p.

Diebold F. and Canlin Li (2006). Forecasting the term structure of government bond yields. Journal of Econometrics. 130:337-364.

Di Matteo T., Aste T., Hyde S.T. and Ramsden S. (2005). Interest rates hierarchical structure. Physica A 355:21-33.

Haykin S. (1999). Neural Networks. A Comprehensive Foundation. Prentice-Hall International. London, 842 p.